2012

The population structure of Escherichia coli isolated from subtropical and temperate soils

Muruleedhara N. Byappanahalli
United States Geological Survey

Tao Yan
University of Hawaii at Manoa

Matthew J. Hamilton
University of Minnesota

Satoshi Ishii
University of Minnesota

Roger S. Fujioka
University of Hawaii

See next page for additional authors

Science of the Total Environment 417-418 (2012) 273–279

Contents lists available at SciVerse ScienceDirect

Science of the Total Environment

journal homepage: www.elsevier.com/locate/scitotenv

The population structure of *Escherichia coli* isolated from subtropical and temperate soils

Muruleedhara N. Byappanahalli [a], Tao Yan [b], Matthew J. Hamilton [c], Satoshi Ishii [c,d], Roger S. Fujioka [e], Richard L. Whitman [a], Michael J. Sadowsky [c,f,*]

[a] *Lake Michigan Ecological Research Station, United States Geological Survey, Porter, IN 46304, USA*
[b] *Department of Civil and Environmental Engineering, Holmes Hall, University of Hawaii at Manoa, Honolulu, HI 96822, USA*
[c] *Department of Soil, Water & Climate, University of Minnesota, St. Paul, MN 55108, USA*
[d] *Division of Environmental Engineering, Hokkaido University, Sapporo, Hokkaido 060-8628, Japan*
[e] *Water Resources Research Center, University of Hawaii, Honolulu, HI 96822, USA*
[f] *BioTechnology Institute, University of Minnesota, St. Paul, MN 55108, USA*

ARTICLE INFO

Article history:
Received 28 July 2011
Received in revised form 6 December 2011
Accepted 20 December 2011
Available online 20 January 2012

Keywords:
Autochthonous population
Indicator bacteria
Biomes
DNA fingerprints
Population structure
Soil environment

ABSTRACT

While genotypically-distinct naturalized *Escherichia coli* strains have been shown to occur in riparian soils of Lake Michigan and Lake Superior watersheds, comparative analyses of *E. coli* populations in diverse soils across a range of geographic and climatic conditions have not been investigated. The main objectives of this study were to: (a) examine the population structure and genetic relatedness of *E. coli* isolates collected from different soil types on a tropical island (Hawaii), and (b) determine if *E. coli* populations from Hawaii and temperate soils (Indiana, Minnesota) shared similar genotypes that may be reflective of biome-related soil conditions. DNA fingerprint and multivariate statistical analyses were used to examine the population structure and genotypic characteristics of the *E. coli* isolates. About 33% (98 of 293) of the *E. coli* from different soil types and locations on the island of Oahu, Hawaii, had unique DNA fingerprints, indicating that these bacteria were relatively diverse; the Shannon diversity index for the population was 4.03. Nearly 60% (171 of 293) of the *E. coli* isolates from Hawaii clustered into two major groups and the rest, with two or more isolates, fell into one of 22 smaller groups, or individual lineages. Multivariate analysis of variance of 89, 21, and 106 unique *E. coli* DNA fingerprints for Hawaii, Indiana, and Minnesota soils, respectively, showed that isolates formed tight cohesive groups, clustering mainly by location. However, there were several instances of clonal isolates being shared between geographically different locations. Thus, while nearly identical *E. coli* strains were shared between disparate climatologically- and geographically-distinct locations, a vast majority of the soil *E. coli* strains were genotypically diverse and were likely derived from separate lineages. This supports the hypothesis that these bacteria are not unique and multiple genotypes can readily adapt to become part of the soil autochthonous microflora.

1. Introduction

The extraintestinal occurrence of fecal indicator bacteria, including *Escherichia coli* and enterococci, has been documented since the early 1960s (Geldreich et al., 1962, 1964; Mundt, 1962) and is now widely accepted by the scientific community and public health officials (Sadowsky and Whitman, 2010). Much of the early work on soil as a potential source of *E. coli* and enterococci came from studies done in tropical biomes, in particular Hawaii (Hardina and Fujioka, 1991) and Guam (Fujioka et al., 1999) where these bacteria were commonly found in stream/river bank soils. Subsequently, more detailed investigations done on the island of Oahu, Hawaii indicated that *E. coli* and enterococci were common in a range of soils, with bacterial densities ranging from <1 to over 10^4/g soil (Byappanahalli, 2000). *E. coli* has similarly been recovered in other tropical environments, including pristine forest soils of Puerto Rico (Lasalde et al., 2005) and from river bank soils in south Florida (Desmarais et al., 2002; Solo-Gabriele et al., 2000).

Recent investigations have also shown that *E. coli* is widely distributed in soils beyond those initially investigated in tropical/subtropical climates. For example, *E. coli* has been recovered in relatively undisturbed, riparian soils of southern Lake Michigan (Byappanahalli et al., 2003, 2006) and in several organic and loamy soils of the Lake Superior watershed (Ishii et al., 2006a). Byappanahalli et al. (2006) recovered *E. coli* in 88% of the soil samples collected over an 8-month period from six exclosure plots along undisturbed forest soils of a Lake Michigan watershed, with bacterial densities ranging from <1 to 1657 MPN/g soil. *E. coli* was similarly recovered in the soils of

* Corresponding author at: Department of Soil, Water, and Climate, 1991 Upper Buford Circle, 439 Borlaug Hall, St. Paul, MN 55108, USA. Tel.: +1 612 626 0977.
E-mail address: Sadowsky@umn.edu (M.J. Sadowsky).

0048-9697/$ – see front matter © 2012 Elsevier B.V. All rights reserved.
doi:10.1016/j.scitotenv.2011.12.041

Lake Superior watersheds (near Duluth, Minnesota), with the highest densities (3.0×10^3 CFU/g soil) occurring between summer and early fall (June to October) and the lowest densities (1.0 CFU/g soil) in the winter (Ishii et al., 2006a).

In addition to soil, *E. coli* has also been recovered from other substrates, including river and lake sediments (Gary and Adams, 1985; Gerba and McLeod, 1976; Obiri-Danso and Jones, 2000) and beach sand (Ghinsberg et al., 1994; Mendes et al., 1997; Whitman and Nevers, 2003). While the environmental occurrence of *E. coli* has been postulated to be due to an adaptive process leading to the formation of naturalized populations (Ishii and Sadowsky, 2008; Winfield and Groisman, 2003), the mechanism by which this process occurs remains unclear.

The HFERP rep-PCR DNA fingerprinting technique (Johnson et al., 2004) has been used to examine the population structure and genetic relatedness of soil *E. coli* isolates obtained from Lake Michigan and Lake Superior watersheds (Byappanahalli et al., 2006; Ishii et al., 2006a). Two significant findings of these studies were that: (a) while some *E. coli* isolates with identical rep-PCR DNA fingerprints were often found in the studied soils, not all *E. coli* isolates were genetically identical; and (b) the soil *E. coli* isolates examined were genetically-distinct from those derived from wildlife, such as deer, geese, seagulls, and terns, inhabiting the watersheds examined. Similarly, using a combination of PCR (for *uidA* gene) and denaturing gradient gel electrophoresis (DGGE) techniques, Lasalde et al. (2005) found genetically heterogeneous *E. coli* populations in soils in both pristine forest and urbanized areas of Puerto Rico. These findings support the hypothesis that naturalized *E. coli* populations occur in disparate and geographically-distinct environments. However, the original source(s) of soil-borne *E. coli* remains unknown at this time.

Based on the findings that metabolically diverse *E. coli* populations occur in soils on Oahu, Hawaii, Fujioka and Byappanahalli (2001) proposed that soil environments likely do not select for a few dominant *E. coli* strains, but rather that numerous strains have the ability to colonize, persist, and adapt to the soil habitat and become part of the resident microflora. Our recent findings showing the occurrence of genotypically-distinct, yet diverse *E. coli* populations in riparian soils of Lake Michigan and Lake Superior watersheds (Byappanahalli et al., 2006; Ishii et al., 2006a) support this hypothesis. Likewise, genotypically diverse *E. coli* strains have been found in water and bank soils along four perennial streams in Oahu, Hawaii (Goto and Yan, 2011). Collectively, these results show that genetically-diverse *E. coli* populations exist in both tropical and temperate soils. However, a broader understanding of the population structure is needed to determine whether: (a) genotypically similar *E. coli* strains exist in diverse soils across a range of geographic and climatic conditions, and (b) *E. coli* genotypes in soil are similar to those reported in other natural habitats, such as beach water freshwater and sediments (Luo et al., 2011). Recently, it has been suggested that *E. coli* genomic composition contributes to its occurrence and persistence in different niches (van Elsas et al., 2010); additional studies are needed to better elucidate these ecological processes.

The objectives of the current study were to: (1) examine the population structure and genetic characteristics of *E. coli* population collected from different soil types and locations from a tropical island (Hawaii) using the rep-PCR DNA fingerprinting technique; and (2) determine if the population structure of *E. coli* from tropical (Hawaii) and temperate (Indiana, Minnesota) soils were disparate or contained shared genotypes that may be reflective of specific soil *E. coli* strains.

2. Materials and methods

2.1. Sampling locations

Soil samples were collected from two main study locations: the island of Oahu, Hawaii (HI; tropical biome) and the riparian soils of southern Lake Michigan (Indiana — IN) and western Lake Superior (Minnesota — MN) watersheds (both temperate biomes). For HI, samples of surface soil (0–6 cm depth) representing the seven major soil associations (USDA Soil Conservation Service and the University of Hawaii Agricultural Experiment Station, 1972) were collected on the island of Oahu between 1991 and1997. The general characteristics of these soil associations and the sampling location names are provided elsewhere (Byappanahalli, 2000). Soil samples were also collected at six different locations along the Manoa Stream network between November 2009 and January 2010. At each location, samples were collected at the stream edge and 1, 3, and 5 m landward from the stream edge. The sites and sampling procedures used for IN and MN locations are provided elsewhere (Byappanahalli et al., 2006; Ishii et al., 2006a). Briefly, soil samples (IN) were collected between March and October 2003 from the Dunes Creek watershed, located within the boundaries of Indiana Dunes State Park and Indiana Dunes National Lakeshore, along the shore of southern Lake Michigan in northwest Indiana. The sampling sites were dominated by white oak and were well above the normal flood zone of Dunes Creek (Byappanahalli et al., 2006).

For MN, the three sampling sites were primarily located in the Lake Superior watershed near Duluth, MN and included the Kingsbury Stark (KS), Nemadji Weinstein (NW), and the St. Louis Clyde (SC) sites (Ishii et al., 2006a). Soil samples were collected between October 2003 and October 2004. The KS site is adjacent to the intersection between Kingsbury Creek and Stark Avenue in Proctor, MN.

2.2. Microbiological analyses

Soil samples (HI) were analyzed for *E. coli* by using two different methods. Bacteria were elutriated from soils using phosphate-buffered water, PBW (pH 7.0 ± 0.2) (APHA, 1989) or ammonium phosphate buffer. Elutriates were diluted as necessary and analyzed for *E. coli* by using the most probable number (MPN) method (soil samples from 1991 to 1997) (APHA, 1989) or by using the membrane filtration method and modified mTEC agar (2009–2010 samples) (APHA, 2005), after pre-filtration of the elutriates through GN-6 membranes (Pall Life Science; Port Washington, NY).

Presumptive *E. coli* cultures from the MPN tubes were confirmed by their growth and colony color on eosine methylene blue (EMB) agar plates (i.e., blue-black colonies with dark centers and green metallic sheen), by production of β-glucuronidase activity, and by speciation by using the API 20 E identification system (bioMerieux Vitek, Inc., Hazelwood, MO). Many of the *E. coli* isolates from this collection were also confirmed by using Biolog GN plates (Fujioka and Byappanahalli, 2001) and by IMViC analyses (APHA, 2005) of randomly-selected *E. coli* colonies from modified mTEC agar plates. All confirmed *E. coli* isolates were stored at −80 °C until used.

Detailed analytical procedures for the isolation of *E. coli* from MN soils are described elsewhere (Ishii et al., 2006a). Briefly, aliquots (10–20 g) of fresh soil were extracted with 0.1 M gelatin-ammonium phosphate extraction solution and the soil-buffer mixture was filtered through 0.45-μm membrane filters (Millipore, Billerica, MA). Membrane filters were transferred to mFC agar medium (Difco, Detroit, MI) and incubated at 35 °C for 2 h, followed by 44.5 °C for 22 h. Well-isolated, blue colonies appearing from mFC plates were streaked onto on mFC agar and presumptive *E. coli* colonies were spot inoculated onto MacConkey (Difco) and CHROMagar ECC (CHROMagar Microbiology, Paris, France). Typical *E. coli* colonies were verified by IMViC and other biochemical analyses, and isolated colonies were stored at −70 °C until used.

The various tests used for confirming *E. coli* isolates from the IN location are provided elsewhere (Byappanahalli et al., 2006). Briefly, soil samples were analyzed for *E. coli* by using the Colilert-18 (Quanti-Tray 2000) method (IDEXX, West Brook, ME). Bacterial growth from fluorescing Quanti-Tray wells were streaked on mTEC

agar plates were incubated at 44.5 °C for 20–22 h; yellow or yellow-brown colonies were then confirmed as *E. coli* by substrate test (urease activity) (APHA, 1998). Well-isolated colonies were further sub-cultured on MacConkey agar (Difco); pink colonies were later confirmed for β-glucuronidase activity on nutrient agar (Difco) containing 4-methylumbelliferyl-β-D-glucuronide, MUG. Confirmed *E. coli* isolates were stored at −80 °C in tryptic soy broth containing 10% glycerol until used.

2.3. Horizontal, fluorophore-enhanced rep-PCR DNA fingerprinting (HFERP)

In the current study, only the HI-1 and HI-2 *E. coli* isolates (100 and 193, respectively; total of 293 isolates) were fingerprinted. For the *E. coli* isolates from IN and MN collections, we used DNA fingerprints that were previously established and stored in our DNA fingerprint library as image files. The DNA fingerprints of 293 *E. coli* isolates collected from the Hawaiian soils (*HI-1*:100 and *HI-2*:193) were obtained using 6 carboxyfluorescein (6-FAM)-labeled Box A1R primers and the HFERP technique as previously described (Johnson et al., 2004). Since then, this technique has been shown to be highly reliable and reproducible in a number of studies (Byappanahalli et al., 2006; Ishii et al., 2006a, 2006b).

HFERP Gel images were analyzed by using BioNumerics v. 2.5 software (Applied-Maths, Sint-Martens-Latem, Belgium) and normalized using the GeneScan-2500 ROX internal lane standard (Applied Biosystems, Foster City, CA). Similarity among DNA fingerprint data was calculated using the curve-based, Pearson coefficient, and dendrograms were generated by using the unweighted pair-group method using arithmetic averages (UPGMA). Clustering of isolates was accomplished by multivariate analysis of variance (MANOVA), a form of discriminants analysis accounting for variance (Dombek et al., 2000), and by principal component analysis (PCA). The *E. coli* population structure and genetic relatedness of IN and MN soil isolates have been described elsewhere (Byappanahalli et al., 2006; Ishii et al., 2006a). In this current study, unique DNA fingerprints from 21 and 106 *E. coli* strains originating from IN and MN soils, respectively, were used for the comparison purposes. As previously defined elsewhere, *E. coli* strains having HFERP DNA fingerprints that were ≥92% similar were considered to be genotypically identical (Ishii et al., 2006a; Johnson et al., 2004). The repeatability of the rep-PCR and HFERP DNA fingerprint methods was confirmed by including a reference *E. coli* strain (isolate 294 from pig) on each gel (Johnson et al., 2004).

The Shannon diversity index (H') was used to calculate the genetic diversity of the soil-borne *E. coli* strains. The diversity index was calculated as follows:

$$H' = -\sum_{i=1}^{S} p_i \ln p_i$$

where S is the number of unique genotypes; p_i is the number of isolates sharing the same genotype [i] over the total number of isolates (Magurran, 1988).

3. Results and discussion

3.1. Isolation of E. coli

E. coli was recovered in most of the samples analyzed, with high variation, both within and between locations: Hawaii, 1.21 ± 0.17 Log MPN/g, $n = 57$ (Byappanahalli, 2000); Indiana, 16 MPN/g ranging <1 to 1657, $n = 66$ (Byappanahalli et al., 2006); and Minnesota, <1 to 3×10^3 CFU/g (Ishii et al., 2006a).

All presumptive *E. coli* isolates ($n = 100$) from HI-1 collection were confirmed by the API 20E biochemical tests; presumptive isolates from HI-2 collections ($n = 193$) were confirmed by similar biochemical tests such as IMViC. Presumptive isolates from MN were confirmed as *E. coli* by a series of biochemical tests, including indole and methyl red tests, the inability to grow on citrate agar, and the presence of β-D-glucuronidase activity using EC-MUG broth (Difco). Only isolates showing atypical responses to any of these tests were further examined by API 20E tests (Ishii et al., 2006a).

Additionally, six isolates from IN were further verified as *E. coli*, following PCR amplification of the 1.5-kb region of the 16S rDNA, encompassing nucleotides 27–1525 (*E. coli* MG1655) using primers 27F and 1525R (Lane, 1991). Sequence analysis of the nearly full-length 16S rRNA gene and the sequenced regions from all six isolates had >98% nucleotide identity to the 16S rDNA from *E. coli* strain MG1655 (accession number U00096) (Byappanahalli et al., 2006). These results indicated that the analytical methods used were optimal for isolating and identifying *E. coli* directly from the soils examined. This is similar to what was previously reported by Ishii et al. (2006a).

3.2. Population structure and genetic relatedness of E. coli isolates collected from Hawaiian soils

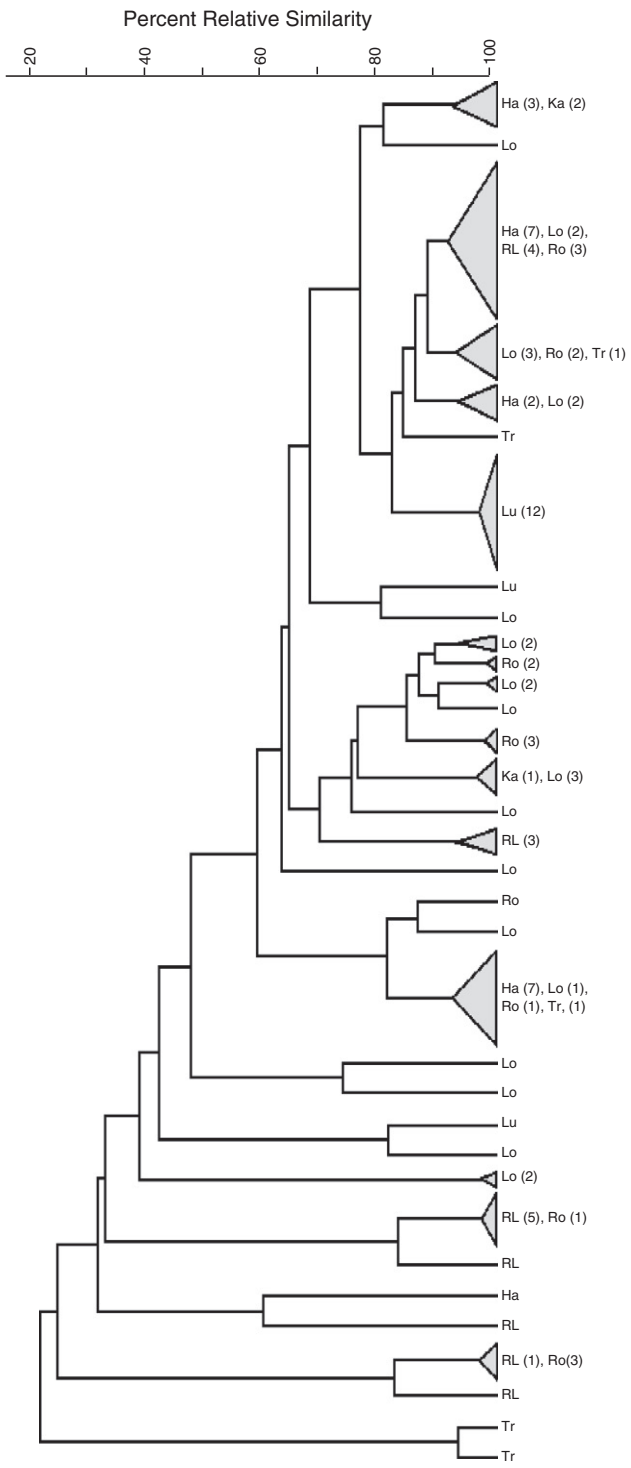
The *E. coli* isolates obtained from different soil types across the Island of Oahu (*HI-1*:100) displayed a high degree of genetic diversity, with no apparent relationship between soils and *E. coli* genotype (Fig. 1). A vast majority of the isolates (81%) clustered into 15 groups, each containing at least two isolates. Several isolates in each soil type were highly related to each other (over 98% similarity), yet they did not cluster by soil type, indicating that there was some degree of diversity in the population within and between soil types.

When we included another 193 isolates (i.e., *HI-2*), which were all collected from a narrower region along a perennial stream (Manoa Stream) for comparison purposes, the soil isolates were again highly diverse and clustered into 216 groups using 92% as the cutoff limit for defining identity. However, because of the large population size, an entire dendrogram could not be presented here. When compressed to reduce overall tree length, 274 (94%) *E. coli* isolates clustered into 24 groups (each containing at least 2 isolates) and the remaining 19 isolates fell into individual lineages (Fig. 2). Nearly 58% of the isolates (171 of 293) clustered into two major groups, with DNA fingerprint similarities ranging from 80 to >99%. Moreover, 18 of the smaller clusters, each containing two or more isolates, comprised genotypes that were related to each other at similarity values ranging from 81 to >99%.

The Shannon diversity index for the *E. coli* population examined was 4.03; Shannon diversity indices for the two sub-populations, *HI-1* (100 isolates) and *HI-2* (193 isolates) were 3.03 and 3.59, respectively. In general, these indices were lower than that recently reported for *E. coli* population associated with *Cladophora* in Lake Michigan (i.e., 5.39, $n = 835$) (Byappanahalli et al., 2007). Such high diversity in soil *E. coli* population may reflect the sources from which they were originally derived, but additional studies are needed to understand such population characteristics not only in soil but, also in a wide range of homoeothermic (e.g., GI tract of humans and warm-blooded animals) and heterothermic (i.e., extraintestinal) habitats.

3.3. Interrelatedness of E. coli obtained from geographically different sites

DNA fingerprint analyses were used to determine the relationship between genetic diversity of *E. coli* and spatial location. For this purpose, we used only unique genotypes (89, 21, and 106 for the HI, IN, and MN locations, respectively). This allowed the study of genetic diversity in the absence of clonality. Of the 216 unique fingerprints



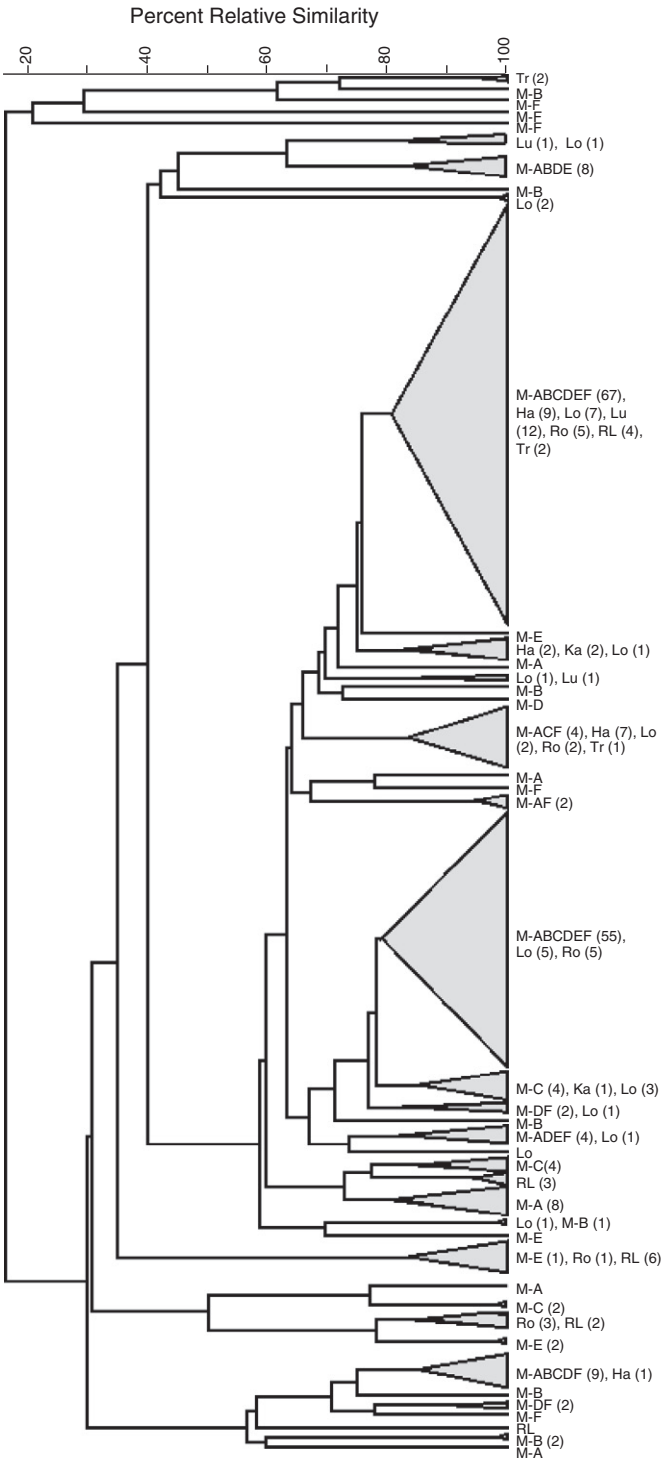


Fig. 1. Dendrogram showing the relative genetic relatedness of *E. coli* isolates (*HI-1*:100) isolated from the 7 major soil associations on the island of Oahu, Hawaii, as determined by HFERP DNA fingerprint analysis using Box A1R primer. DNA fingerprint similarities were calculated by using the curve-based, Pearson coefficient and the dendrogram was generated by using the unweighted pair-group method with arithmetic averages (UPGMA).

Fig. 2. Dendrogram showing the relative genetic relatedness of soil *E. coli* isolates ($n = 293$) collected on the island of Oahu, Hawaii: (a) *HI-1*:100 isolates were collected from samples representing the seven major soil associations found on Oahu and (b) *HI-2*:193 isolates were obtained from stream bank soils at six different locations along the Manoa Stream on Oahu. Due to the large population size examined, a condensed dendrogram is presented here showing 24 groups and 19 individual lineages. The numbers in parentheses adjacent to each cluster indicate the number of isolates contained in that cluster. DNA fingerprint similarities were calculated by using the curve-based, Pearson coefficient and the dendrogram was generated by using the unweighted pair-group method with arithmetic averages (UPGMA).

examined, 162 (75%) clustered into 40 groups, with each group containing 2 or more genotypes (Fig. 3). The remaining 54 genotypes were comprised of individual lineages. Overall, genotypes were clustered by location.

Interestingly, there were several instances where genotypes were shared between and among strains originating from geographically-distinct locations. For example, group 1 contained five strains (HI, 4 and MN, 1) — three of them were 80 to 93% similar to each other; one each from HI and MN were 93% similar. Group 2 comprised 25

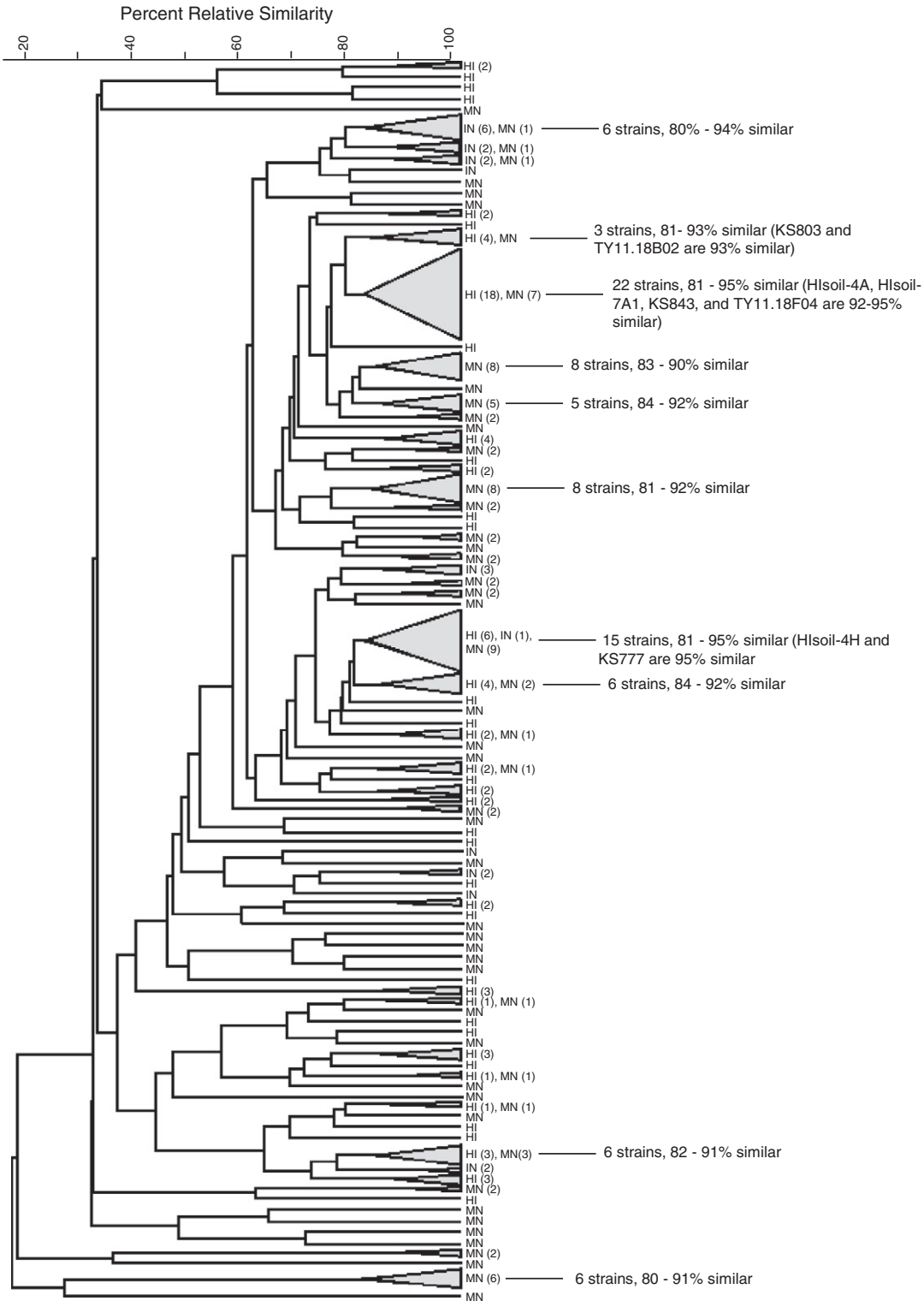


Fig. 3. Dendrogram showing the interrelatedness of soil *E. coli* population from Hawaii (HI), Indiana (IN), and Minnesota (MN) sites. Only unique genotypes found in these soils were used in this analysis (HI, 89, IN, 21, and MN, 106; $n = 216$). The numbers in parentheses adjacent to each cluster indicate the number of isolates contained in that cluster. DNA fingerprint similarities were calculated by using the curve-based, Pearson coefficient and the dendrogram was generated by using the unweighted pair-group method with arithmetic averages (UPGMA).

strains (HI, 18 and MN, 7) and 22 of them had similarity values ranging from 81 to 95% and 4 strains (HI, 3 and MN, 1) that were similar to each other by ≥92%. Group 3 comprised 16 strains (HI, 6, IN, 1, and MN, 9), with 15 of these isolates had similarity values ranging 81 to 95%, and 2 strains (one each from HI and MN) were 95% similar to each other. This suggests that these strains may have shared a common ancestral lineage, evolutionary history, or were derived from the same source (Luo et al., 2011; Walk et al., 2009). Taken together, these findings support the general hypothesis that: (a) the soil environment likely does not select for one or two dominant strains, but instead numerous strains can survive and persist in this environment (Fujioka and Byappanahalli, 2001); and (b) that the

soil-borne *E. coli* strains share certain characteristics that are different from their intestinal counterparts (Luo et al., 2011).

Multivariate analysis of variance was used to examine the clustering of the soil *E. coli* isolates obtained from tropical and temperate biomes (Fig. 4). The first and second discriminants accounted for ~92% of the canonical variation (63% and 28%, respectively). Ninety-two of the HI strains derived from different soil types and locations clustered into two separate groups. Overall the strains in these two groups were generally more closely related to each other than to *E. coli* strains isolated from the IN and MN locations. Nonetheless, a high degree of genetic diversity was apparent within each group, indicating the wide heterogeneity in the population. Likewise, the strains from IN and MN were different from each other and clustered into two distinct groups.

Interestingly, a few of the MN strains were nearly identical to those from HI (e.g., isolates KS804, MN and 3K, HI; 92.4%; KS145, MN and 4O and 4P, HI; 92.2%; KS777, MN and 4F, 4G, and 4H, HI; 94.9%). Similarly, a few IN strains were more closely related to MN strains, and as a population, IN strains were different from HI strains. The broad clustering of *E. coli* by location further supports the hypothesis that soil environments do not select for one or two dominant strains, instead genetically diverse strains, presumably from different sources may have colonized these soils over time. At the same time, heterogeneous soil edaphic factors (e.g., desiccation, nutrient status, or competition) may also contribute to high strain diversity. Furthermore, results of this study indicate that the soils examined have the capacity to support the persistence of genetically diverse *E. coli* strains, as has been reported in other studies (Bergholz et al., 2011; Brennan et al., 2010; Byappanahalli et al., 2006; Goto and Yan, 2011; Ishii et al., 2006a; Texier et al., 2008), and potentially their growth under certain conditions (Byappanahalli and Fujioka, 2004; Ishii et al., 2006a; Whitman et al., 2006).

The genetic processes leading to the occurrence of *E. coli* in soils remain speculative, but some of the mechanisms may include selection for strains capable of surviving and growing in these habitats (i.e., naturalized populations) (Ishii et al., 2006a; Texier et al., 2008; Topp et al., 2003) and the ability to use different carbon and energy sources (i.e., nutritional/metabolic diversity) (Fujioka and Byappanahalli, 2001). In addition, a variety of physical and biological factors, such as available moisture (Byappanahalli et al., 2003; Chandler and Craven, 1980; Solo-Gabriele et al., 2000) and nutrients (Byappanahalli and Fujioka, 2004; Ishii et al., 2010; Recorbet et al., 1992), predation (Solo-Gabriele et al., 2000; Sorensen et al., 1999), and other edaphic factors (Bergholz et al., 2011; Ishii et al., 2010) likely contribute to *E. coli* survival and persistence in soil environments.

Aside from soil, naturalized *E. coli* populations have also been found in lake and river watersheds (Kon et al., 2009; Power et al., 2005), beach sand (Ishii et al., 2007; Kon et al., 2009), and aquatic vegetation, such as green algae *Cladophora* (Badgley et al., 2010; Byappanahalli et al., 2007). Thus, these *E. coli* populations represent clades that are clearly distinct from *E. coli* found in intestinal environments (Luo et al., 2011).

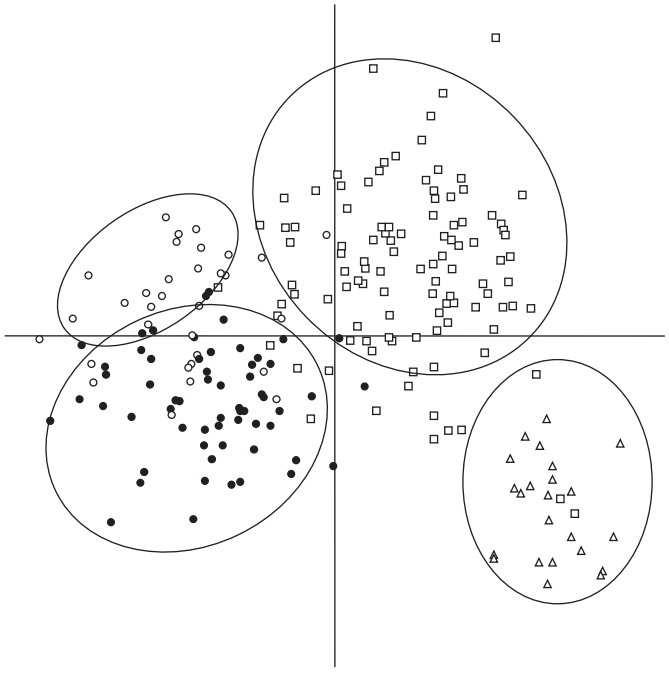


Fig. 4. Multivariate analysis of variance (MANOVA) of HFERP DNA fingerprints of unique *E. coli* genotypes obtained from tropical (HI) and temperate (IN and MN) soil biomes. A binary band-matching character table was analyzed by MANOVA, accounting for variance. Only the first two discriminants are presented in this figure and the first two dimensions account for 91% of the canonical variation among the data. Figure legend for unique soil *E. coli* genotypes: HI isolates (●) and (○); IN isolates (△); and MN isolates (□).

4. Summary and conclusions

Previously, studies have attempted to understand the population structure of *E. coli* in soil environments; such studies, however, have been mostly site-specific. For instance, Byappanahalli et al. (2006) and Ishii et al. (2006a, 2006b) examined *E. coli* populations in soils from Lake Michigan and Lake Superior watersheds, respectively. In both these studies, soil *E. coli* populations comprised genetically heterogeneous strains, yet as a population, the soil strains were clearly distinct from *E. coli* strains derived from wildlife (gulls, terns, deer and most geese) commonly found in these watersheds. Along these lines, Lasalde et al. (2005) found heterogeneous *E. coli* populations in pristine tropical forest soils. Collectively, these data suggest that genotypically-distinct yet diverse *E. coli* populations occur in soil across biomes.

Despite increasing evidence that *E. coli* has a niche in soil environments, many of the ecological questions relating to its original source, colonization potential, and the nature of its interactions with other microflora remain largely speculative at this time. Our working hypothesis, however, is that the soil *E. coli* strains may have originated from fecal deposits of animals or other sources, and over time became established as part of the normal soil microbiota.

In conclusion, unique *E. coli* strains were found in both temperate and tropical soils; nonetheless, there was a high degree of genetic diversity within the population in both of these soil biomes. There was some evidence of strains being shared between locations, however, as a population, the soil-borne *E. coli* strains formed a tight, cohesive genetic group clustering by location. These findings support the hypothesis that *E. coli* strains associated with soils represent a unique group of bacteria. The presence of naturalized *E. coli* in tropical and temperate soils, even in frozen temperate soils, will likely confound the use of *E. coli* as a reliable indicator of fecal contamination. Thus, soil should be considered as an environmental source of indicator bacteria, potentially influencing shoreline water quality.

Acknowledgments

This work was supported, in part, by grants from the University of Minnesota Agricultural Experiment Station and the BioTechnology Institute (to M.J.S.), by grants from the Minnesota Sea Grant program (to M.J.S.), and by training grant 2T32-GM008347 from the National Institutes of Health (to M.J.S.). Any use of trade, product or firm names is for descriptive purposes only and does not imply endorsement by the US Government. This article is Contribution 1677 of the USGS Great Lakes Science Center.

References

APHA. Standard methods for the examination of water and wastewater. 17th ed. Washington, D.C.: American Public Health Association; 1989

APHA. Standard methods for the examination of water and wastewater. 20th ed. Washington, D.C.: American Public Health Association; 1998

APHA. Microbiological examination: part 9000. Standard methods for the examination of water and wastewater. 21st ed. Washington, D.C.: American Public Health Association; 2005

Badgley BD, Ferguson J, Kleinheinz GT, McDermott CM, Sandrin TR, Kinzelman J, et al. Multi-scale temporal and spatial variation in genotypic composition of *Cladophora*-borne *Escherichia coli* populations in Lake Michigan. Water Res 2010;45:721–31.

Bergholz PW, Noar JD, Buckley DH. Environmental patterns are imposed on the population structure of *Escherichia coli* after fecal deposition. Appl Environ Microbiol 2011;77:211–9.

Brennan FP, O'Flaherty V, Kramers G, Grant J, Richards KG. Long-term persistence and leaching of *Escherichia coli* in temperate maritime soils. Appl Environ Microbiol 2010;76:1449–55.

Byappanahalli MN. Assessing the persistence and multiplication of fecal indicator bacteria in Hawaii's soil environment. [dissertation], Department of Microbiology, Honolulu (HI): University of Hawaii at Manoa, p. 244; 2000.

Byappanahalli M, Fujioka R. Indigenous soil bacteria and low moisture may limit but allow faecal bacteria to multiply and become a minor population in tropical soils. Water Sci Technol 2004;50:27–32.

Byappanahalli M, Fowler M, Shively D, Whitman R. Ubiquity and persistence of *Escherichia coli* in a midwestern coastal stream. Appl Environ Microbiol 2003;69:4549–55.

Byappanahalli MN, Whitman RL, Shively DA, Sadowsky MJ, Ishii S. Population structure, persistence, and seasonality of autochthonous *Escherichia coli* in temperate, coastal forest soil from a Great Lakes watershed. Environ Microbiol 2006;8:504–13.

Byappanahalli MN, Whitman RL, Shively DA, Ferguson J, Ishii S, Sadowsky M. Population structure of *Cladophora*-borne *Escherichia coli* in nearshore water of Lake Michigan. Water Res 2007;41:3649–54.

Chandler DS, Craven JA. Relationship of soil moisture to survival of *Escherichia coli* and *Salmonella typhimurium* in soils. Aust J Agr Res 1980;31:547–55.

Desmarais TR, Solo-Gabrielle HM, Palmer CJ. Influence of soil on fecal indicator organisms in a tidally influenced subtropical environment. Appl Environ Microbiol 2002;68:1165–72.

Dombek PE, Johnson LK, Zimmerley ST, Sadowsky MJ. Use of repetitive DNA sequences and the PCR to differentiate *Escherichia coli* isolates from human and animal sources. Appl Environ Microbiol 2000;66:2572–7.

Fujioka RS, Byappanahalli MN. Microbial ecology controls the establishment of fecal bacteria in tropical soil environment. In: Matsuo T, Hanaki K, Takizawa S, Satoh H, editors. Advances in Water and Wastewater Treatment Technology: Molecular Technology, Nutrient Removal, Sludge Reduction and Environmental Health. Amsterdam, the Netherlands: Elsevier; 2001. p. 273–83.

Fujioka R, Sian-Denton C, Borja M, Castro J, Morphew K. Soil: the environmental source of *Escherichia coli* and enterococci in Guam's streams. J Appl Microbiol Symp Suppl 1999;85:83S–9S.

Gary HL, Adams JC. Indicator bacteria in water and stream sediments near the snowy range in southern Wyoming. Water Air Soil Pollut 1985;25:133–44.

Geldreich EE, Huff CB, Bordner RH, Kabler PW, Clark HF. The faecal coli-aerogenes flora of soils from various geographical areas. J Appl Bacteriol 1962;25:87–93.

Geldreich EE, Kenner BA, Kabler PW. Occurence of coliforms, fecal coliforms, and streptococci on vegetation and insects. Appl Microbiol 1964;12:63–9.

Gerba CP, McLeod JS. Effect of sediments on the survival of *Escherichia coli* in marine waters. Appl Environ Microbiol 1976;32:114–20.

Ghinsberg RC, Bar Dov L, Rogol M, Sheinberg Y, Nitzan Y. Monitoring of selected bacteria and fungi in sand and sea water along the Tel Aviv coast. Microbios 1994;77:29–40.

Goto DK, Yan T. Genotypic diversity of *Escherichia coli* in the water and soil of tropical watersheds in Hawaii. Appl Environ Microbiol 2011;77:3988–97.

Hardina CM, Fujioka RS. Soil: the environmental source of *Escherichia coli* and enterococci in Hawaii's streams. Environ Toxicol Water Qual 1991;6:185–95.

Ishii S, Sadowsky MJ. *Escherichia coli* in the environment: implications for water quality and human health. Microbes Environ 2008;23:101–8.

Ishii S, Ksoll WB, Hicks RE, Sadowsky MJ. Presence and growth of naturalized *Escherichia coli* in temperate soils from Lake Superior watersheds. Appl Environ Microbiol 2006a;72:612–21.

Ishii S, Yan T, Shively DA, Byappanahalli MN, Whitman RL, Sadowsky MJ. *Cladophora* (Chlorophyta) spp. harbor human bacterial pathogens in nearshore water of Lake Michigan. Appl Environ Microbiol 2006b;72:4545–53.

Ishii S, Hansen DL, Hicks RE, Sadowsky MJ. Beach sand and sediments are temporal sinks and sources of *Escherichia coli* in Lake Superior. Environ Sci Technol 2007;41:2203–9.

Ishii S, Yan T, Vu H, Hansen DL, Hicks RE, Sadowsky MJ. Factors controlling long-term survival and growth of naturalized *Escherichia coli* populations in temperate field soils. Microbes Environ 2010;25:8-14.

Johnson L-AK, Brown MB, Carruthers EA, Ferguson JA, Dombek PE, Sadowsky MJ. Sample size, library composition, and genotypic diversity among natural populations of *Escherichia coli* from different animals influence accuracy of determining sources of fecal pollution. Appl Environ Microbiol 2004;70:4478–85.

Kon T, Weir SC, Howell T, Lee H, Trevors JT. Repetitive element (REP)-polymerase chain reaction (PCR) analysis of *Escherichia coli* isolates from recreational waters of southeastern Lake Huron. Can J Microbiol 2009;55:269–76.

Lane DJ. 16S/23S rRNA sequencing. In: Stackbrandt E, Goodfellow M, editors. Nucleic Acid Techniques in Bacterial Systematics. Cambridge, UK: John Wiley and Sons Ltd.; 1991. p. 115–75.

Lasalde C, Rodriguez R, Toranzos GA, Smith HH. Heterogeneity of uidA gene in environmental *Escherichia coli* populations. J Water Health 2005;3:297–304.

Luo C, Walk ST, Gordon DM, Feldgarden M, Tiedje JM, Konstantinidis KT. Genome sequencing of environmental *Escherichia coli* expands understanding of the ecology and speciation of the model bacterial species. Proc Natl Acad Sci U S A 2011;108:7200–5.

Magurran AE. Ecological diversity and its measurement. Princeton, USA: Princeton University Press; 1988.

Mendes B, Urbano P, Alves C, Lapa N, Morais J, Nascimento J, et al. Sanitary quality of sands from beaches of Azores Islands. Water Sci Technol 1997;35:147–50.

Mundt JO. Occurrence of enterococci on plants in a wild environment. Appl Microbiol 1962;11:141–4.

Obiri-Danso K, Jones K. Intertidal sediments as reservoirs for hippurate negative campylobacters, salmonellae and faecal indicators in three EU recognised bathing waters in north west England. Water Res 2000;34:519–27.

Power ML, Littlefield-Wyer J, Gordon DM, Veal DA, Slade MB. Phenotypic and genotypic characterization of encapsulated *Escherichia coli* isolated from blooms in two Australian lakes. Environ Microbiol 2005;7:631–40.

Recorbet G, Steinberg C, Faurie G. Survival in soil of genetically engineered *Escherichia coli* as related to inoculum density, predation and competition. FEMS Microbiol Ecol 1992;101:251–60.

Sadowsky MJ, Whitman RL. The fecal bacteria. In: Sadowsky MJ, Whitman RL, editors. Washington, D.C.: ASM Press; 2010

Solo-Gabriele HM, Wolfert MA, Desmarais TR, Palmer CJ. Sources of *Escherichia coli* in a coastal subtropical environment. Appl Environ Microbiol 2000;66:230–7.

Sorensen SJ, Schyberg T, Ronn R. Predation by protozoa on *Escherichia coli* K12 in soil and transfer of resistance plasmid RP4 to indigenous bacteria in soil. Appl Soil Ecol 1999;11:79–90.

Texier S, Prigent-Combaret C, Gourdon MH, Poirier MA, Faivre P, Dorioz JM, et al. Persistence of culturable *Escherichia coli* fecal contaminants in dairy alpine grassland soils. J Environ Qual 2008;37:2299–310.

Topp E, Welsh M, C. TY, Dang A, Lazarovits G, Conn K, et al. Strain-dependent variability in growth and survival of *Escherichia coli* in agricultural soil. FEMS Microbiol Ecol 2003;44:303–8.

USDA. Soil survey of islands of Kauai, Oahu, Maui, Molokai, and Lanai, State of Hawaii. Washington, D.C.: Soil Conservation Service, United States Department of Agriculture in cooperation with the University of Hawaii Agricultural Experiment Station; 1972.

van Elsas JD, Semenov AV, Costa R, Trevors JT. Survival of *Escherichia coli* in the environment: fundamental and public health aspects. ISME J 2010:1-11.

Walk ST, Alm EW, Gordon DM, Ram RJ, Toranzos GA, Tiedje JM, et al. Cryptic lineages of the genus *Escherichia*. Appl Environ Microbiol 2009;75:6534–44.

Whitman RL, Nevers MB. Foreshore sand as a source of *Escherichia coli* in nearshore water of a Lake Michigan beach. Appl Environ Microbiol 2003;69:5555–62.

Whitman RL, Nevers MB, Byappanahalli MN. Examination of the watershed-wide distribution of *Escherichia coli* along southern Lake Michigan: an integrated approach. Appl Environ Microbiol 2006;72:7301–10.

Winfield MD, Groisman EA. Role of nonhost environments in the lifestyles of *Salmonella* and *Escherichia coli*. Appl Environ Microbiol 2003;69:3687–94.